

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**





**Europäisches  
Patentamt**

**European  
Patent Office**

**Office européen  
des brevets**

**Bescheinigung**

**Certificate**

**Attestation**

Die angehefteten Unterlagen stimmen mit der ursprünglich eingereichten Fassung der auf dem nächsten Blatt bezeichneten europäischen Patentanmeldung überein.

The attached documents are exact copies of the European patent application described on the following page, as originally filed.

Les documents fixés à cette attestation sont conformes à la version initialement déposée de la demande de brevet européen spécifiée à la page suivante.

**Patentanmeldung Nr. Patent application No. Demande de brevet n°**

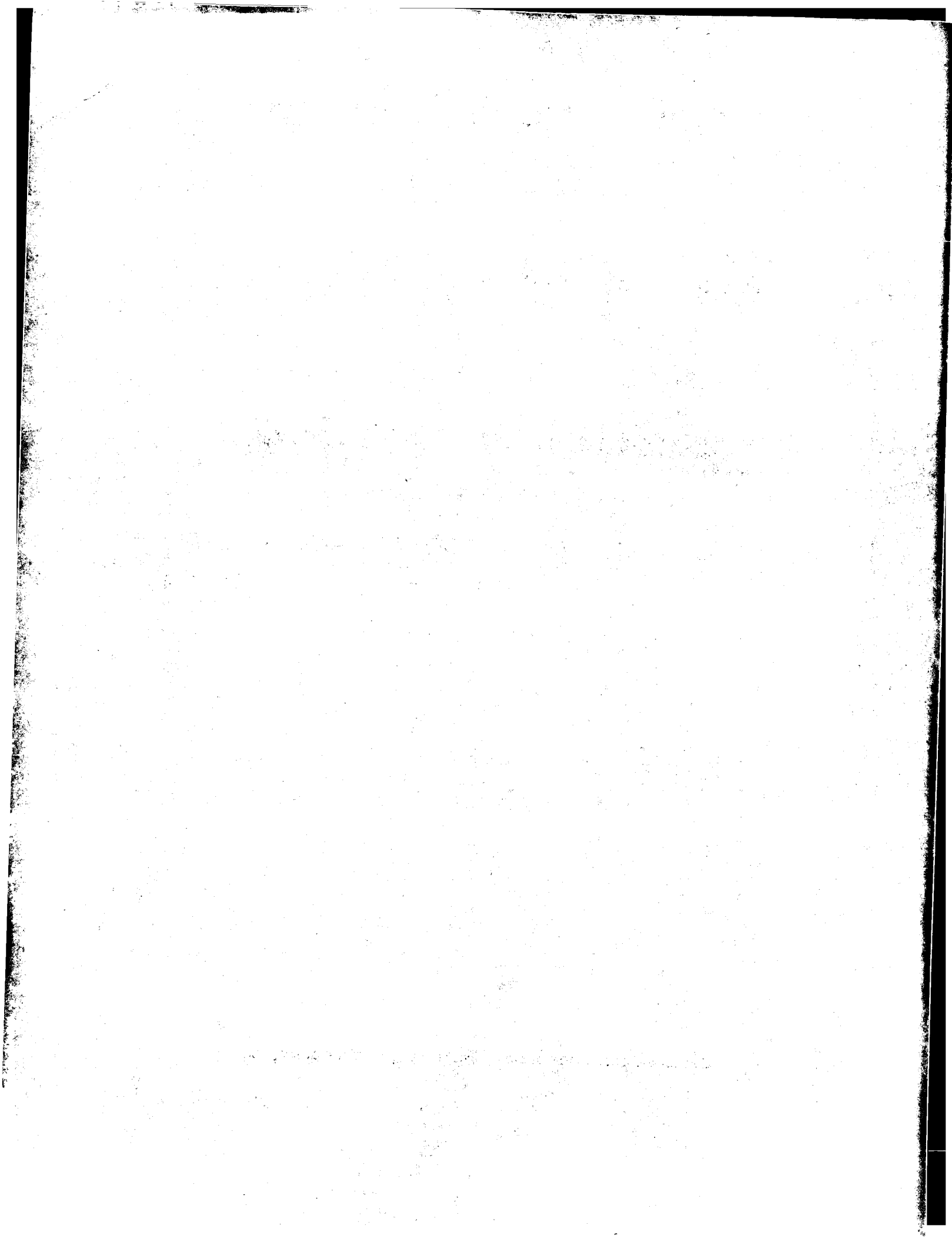
02425601.8

Der Präsident des Europäischen Patentamts;  
Im Auftrag

For the President of the European Patent Office

Le Président de l'Office européen des brevets  
p.o.

**R C van Dijk**





Europäisches  
Patentamt

European  
Patent Office

Office européen  
des brevets

**Blatt 2 der Bescheinigung**  
**Sheet 2 of the certificate**  
**Page 2 de l'attestation**

Anmeldung Nr.:  
Application no.:  
Demande n°: 02425601.8

Anmeldetag:  
Date of filing:  
Date de dépôt: 04/10/02

Anmelder:  
Applicant(s):  
Demandeur(s):  
STMicroelectronics S.r.l.  
20041 Agrate Brianza (Milano)  
ITALY

Bezeichnung der Erfindung:  
Title of the invention:  
Titre de l'invention:

Method and system for processing signals via perceptive vectorial quantization, computer program product therefor

In Anspruch genommene Priorität(en) / Priority(ies) claimed / Priorité(s) revendiquée(s)

Staat:  
State:  
Pays:

Tag:  
Date:  
Date:

Aktenzeichen:  
File no.  
Numéro de dépôt:

Internationale Patentklassifikation:  
International Patent classification:  
Classification internationale des brevets:

H04N7/26, H04N7/30, H04N7/34, H04N7/50

Am Anmeldetag benannte Vertragstaaten:  
Contracting states designated at date of filing:  
Etats contractants désignés lors du dépôt:

AT/BG/BE/CH/CY/CZ/DE/DK/EE/ES/FI/FR/GB/GR/IE/IT/LI/LU/MC/NL/

Bemerkungen:  
Remarks:  
Remarques:

See for the original title of the application, page 1 of the description



"PROCEDIMENTO E SISTEMA PER IL TRATTAMENTO DI SEGNALI  
MEDIANTE QUANTIZZAZIONE VETTORIALE PERCETTIVA, E  
RELATIVO PRODOTTO INFORMATICO

\*\*\*

5 Campo dell'invenzione

La presente invenzione si riferisce alle tecniche per l'elaborazione dei segnali ed è stata sviluppata con particolare attenzione alla possibile applicazione nell'ambito di sistemi che prevedono la riduzione della  
10 quantità di dati richiesta per rappresentare in formato digitale un'immagine (still picture) o una sequenza d'immagini (video).

Descrizione della tecnica nota

Nella tecnica sono conosciute diverse soluzioni  
15 per la compressione efficiente d'immagini digitali. Queste soluzioni sono di solito caratterizzate da un'elevata complessità computazionale e non sono facilmente integrabili nelle soluzioni correntemente denominate SoC (acronimo per System on a Chip).

20 Le tecniche di compressione delle immagini digitali sono classificabili in due gruppi fondamentali.

Un primo gruppo è costituito dalle cosiddette tecniche di compressione senza perdita di qualità  
25 (lossless), adatte anche per il trattamento d'altri dati digitali. Scopo di questo tipo di compressione è rimuovere la ridondanza statistica dei dati.

Ad ogni dato digitale è assegnato un numero variabile di bit che dipende dalla frequenza statistica  
30 di quel particolare dato.

Facendo riferimento, a titolo esemplificativo, alla cosiddetta compressione di Huffman, ad ogni dato digitale è assegnato un numero variabile e intero di bit secondo la seguente regola: ai dati digitali più  
35 frequenti sono assegnati codici binari brevi mentre ai

dati digitali poco frequenti sono assegnati codici binari lunghi.

5 Sono poi note tecniche di compressione aritmetica in cui ad ogni dato digitale è assegnato un numero variabile e frazionario di bit. Il criterio d'assegnazione dei bit è analogo a quello visto per la compressione di Huffman.

10 Altri metodi di compressione sono basati sull'impiego di dizionari. Le sequenze dei dati digitali da comprimere sono ricondotte a parole, di lunghezza variabile, di un dizionario. Ad ogni parola corrisponde un opportuno codice binario di lunghezza fissa o variabile. In questo contesto si inserisce l'algoritmo d'identificazione del dizionario ottimale  
15 dovuto a Lempel e Ziv.

Un secondo gruppo di tecniche di compressione note è costituito dalle tecniche di compressione con perdita di qualità (lossy).

20 Scopo di questo tipo di compressione è di rimuovere la ridondanza percettiva nei dati. L'immagine è modificata eliminando quello che non può essere percepito, o che è percepito meno, dal sistema visivo dell'osservatore (HVS o Human Visual System). La caratteristica più sfruttata del sistema visivo è  
25 riconducibile al fatto che la sensibilità alle basse frequenze è maggiore rispetto a quella alle alte frequenze. Inoltre, la percezione della risoluzione spaziale dell'informazione di luminosità è più marcata rispetto alla percezione dell'informazione di  
30 cromaticità.

La rappresentazione dell'informazione di cromaticità può quindi essere più imprecisa, nel senso che la risoluzione spaziale può essere più bassa. La cromaticità è quindi sottocampionata rispetto alla  
35 luminosità. La perdita di qualità che ne deriva non è



praticamente percepita dall'osservatore.

A titolo d'esempio, per lo standard ITU-R BT.601, il rapporto di sottocampionamento fra il segnale di luminanza (Y) e le due differenze colore (CbCr o UV o IQ o DbDr) è pari a 4:2:2. Per il ben noto standard MPEG il rapporto è di 4:2:0, dove lo 0 indica che il sottocampionamento è sia verticale, sia orizzontale.

In modo analogo, la rappresentazione delle altre sequenze può essere più imprecisa, nel senso di una quantizzazione più rozza, con conseguente risparmio di bit. La perdita di qualità percepita che ne deriva è però bassa a causa della minore sensibilità del sistema visivo a queste frequenze.

La suddivisione in alte e basse frequenze può essere fatta solamente dopo essere passati dal dominio spaziale a quello delle frequenze mediante l'operazione di trasformazione. Le trasformazioni più utilizzate sono, a titolo esemplificativo, la trasformata coseno discreta (DCT o Discrete Cosine Transform) e la trasformata wavelet discreta (DWT o Discrete Wavelet Transform).

#### Scopi e sintesi della presente invenzione

La presente invenzione si prefigge o scopo di fornire una soluzione perfezionata rispetto a quelle della tecnica nota sotto diversi punti di vista ed in particolare per quanto riguarda le esigenze di mantenere bassa la complessità computazionale (cioè il numero d'operazioni d'elaborazione e il numero d'unità di memorizzazione) in vista dell'integrazione della funzione di de/compressione d'immagini in un sistema più complesso. Tutto ciò senza penalizzare in modo marcato l'efficienza di compressione e consentendo quindi la possibilità di ridurre l'area occupata sul silicio dai relativi circuiti, con la conseguente possibilità di ridurre i costi di realizzazione

conseguendo altresì una ridotta dissipazione (fattore, quest'ultimo, particolarmente importante per i dispositivi a bassa potenza o low power).

Secondo la presente invenzione, tale scopo è  
5 conseguito grazie ad un procedimento avente le caratteristiche richiamate in modo specifico nelle rivendicazioni che seguono. L'invenzione riguarda anche il corrispondente sistema, nonché il relativo prodotto informatico direttamente caricabile nella memoria  
10 interna di un elaboratore digitale e contenente porzioni di codice software per realizzare le fasi del procedimento secondo l'invenzione quando il prodotto è eseguito su un elaboratore numerico.

La soluzione secondo l'invenzione è integrabile,  
15 ad esempio, in un'unità digitale per l'acquisizione d'immagini (sensori CCD/CMOS), sia in un'unità di visualizzazione d'immagini (display LCD). In particolare, nel caso di fotocamere digitali e dispositivi analoghi o affini, i dati sono acquisiti  
20 dal sensore secondo una configurazione spaziale o pattern noto come Bayer Color Filter Array (CFA) che permette di associare a ciascun pixel solo una delle tre componenti di colore. La corrispondente immagine RGB è poi ricostruita attraverso una sequenza  
25 d'elaborazione dell'immagine (IGP, acronimo per Image Generation Pipeline) tra le quali è sempre presente un blocco dedicato alla compressione.

Il fatto di anticipare questo passo, collocandolo immediatamente dopo l'acquisizione dei dati dal  
30 sensore, permette di ridurre la banda necessaria per trasmettere l'immagine dall'unità d'elaborazione o memorizzazione al display. Questa soluzione risulta utile soprattutto nel caso d'applicazioni in cui i dati acquisiti debbano essere trasmessi per essere elaborati  
35 in remoto.

In sostanza, la soluzione secondo l'invenzione è basata sulla costruzione di un quantizzatore vettoriale o multidimensionale con celle di quantizzazione non uniformi per array di dati digitali contenenti  
 5 componenti cromatiche (lineari o non lineari).

Il quantizzatore vettoriale è costruito in modo da permettere la contemporanea riduzione della ridondanza statistica e della ridondanza percettiva dei dati contenuti nell'array e in modo da minimizzare la  
 10 complessità della codifica (corrispondente alla compressione) e della decodifica (corrispondente alla decompressione).

#### Breve descrizione dei disegni annessi

L'invenzione sarà ora descritta, a puro titolo  
 15 d'esempio non limitativo, con riferimento ai disegni annessi, nei quali:

- la figura 1 illustra, sotto forma di uno schema a blocchi, lo schema di un circuito codificatore operante secondo l'invenzione,
- 20 - la figura 2 illustra, sempre sotto forma di uno schema a blocchi, la struttura di un circuito decodificatore utilizzabile nell'ambito dell'invenzione,
- la figura 3 illustra lo schema di scansione di una delle componenti cromatiche (in particolare si tratta della componente verde) di un cosiddetto Bayer  
 25 Pattern nell'ambito di un sistema operante secondo l'invenzione, e

- le figure 4 a 7 rappresentano vari esempi d'applicazione della soluzione secondo l'invenzione.

30 Riferendosi dapprima allo schema della figura 1, il blocco indicato con 10 rappresenta in generale un sensore (nel seguito si supporrà che lo stesso sia un sensore d'immagine del tipo correntemente indicato come sensore di Bayer), suscettibile di generare alla sua  
 35 uscita un segnale  $x(i)$  rappresentativo di un'immagine

in forma digitalizzata.

Sensori di questo tipo sono ampiamente noti nella tecnica; le relative caratteristiche di funzionamento, anche per quanto riguarda le caratteristiche dei  
5 segnali generati, non richiedono dunque di essere descritte in modo particolareggiato in questa sede.

Il riferimento 12 indica nella figura 1 un blocco predittore in grado di generare, dall'immagine d'ingresso generata dal blocco 10 un relativo segnale  
10 di predizione  $p(i)$ . Nell'esempio qui illustrato il segnale di predizione è generato, per ciascun segnale  $x(i)$  in funzione del valore precedente  $x(i-1)$ .

Di conseguenza, il segnale  $p(i)$  è esprimibile come:

$$15 \quad p(i) = \text{predictor}(x(i)) = x(i-1).$$

I segnali  $x(i)$  e  $p(i)$  sono poi sommati (con segni opposti) in un modo di somma 12 che genera alla sua uscita un segnale d'errore di predizione  $e(i)$  esprimibile in generale come:

$$20 \quad e(i) = x(i) - p(i) \\ \text{con } e(1) = x(1).$$

Il riferimento 13 indica un blocco in cui il segnale d'errore di predizione  $e(i)$  è sottoposto a quantizzazione per poi essere sottoposto, in un blocco  
25 indicato con 14, ad una funzione di codifica di simbolo che sta alla base dell'immagine d'uscita, indicata complessivamente con 15.

Lo schema di codifica rappresentato - a puro titolo esemplificativo, e dunque senza intenti di  
30 limitazione della portata dell'invenzione, che è affatto generale - nella figura 1 è quindi essenzialmente del tipo correntemente denominato DPCM (o PCM differenziale). Il carattere lossy della codifica è evidentemente legato alla funzione di  
35 quantizzazione attuata nel blocco 13.

La linea a tratti indicata con L indica che, nel caso di una codifica senza perdita di qualità (lossless) gli errori di predizione non sono quantizzati prima della codifica attuata nel blocco 14.

5 Lo schema a blocchi della figura 2 illustra la funzione complementare di decodifica. Qui l'immagine in ingresso 25 (costituita da una immagine compressa - virtualmente identificabile con l'immagine 15 derivante dall'azione di codifica) è alimentata ad un blocco di  
10 (de)codifica di simbolo 26 per poi passare ad un blocco di de(quantizzazione) 27 dove il segnale d'errore di predizione  $e(i)$  è ricostruito per poi essere sommato con segno, in un nodo indicato con 28, ad un segnale di predizione  $p(i)$  generato da un blocco di predizione 29.  
15 Tutto questo per generare un segnale corrispondente all'immagine d'uscita ricostruita, rappresentata dal blocco 30.

Si apprezzerà che anche nel caso del decodificatore DPCM della figura 2 il blocco di  
20 predizione 29 opera in funzione del valore del segnale ricostruito relativo all'istante di tempo precedente.

In altre parole, la funzione decodifica rappresentata nello schema della figura 2 corrisponde all'implementazione della relazione:

$$\begin{aligned} 25 \quad & x(i) = e(i) + p(i) \\ & \text{con } x(1) = e(1). \end{aligned}$$

Anche nello schema del decodificatore della figura 2, la linea a tratti indicata con L illustra il possibile ricorso a tecniche di decodifica senza  
30 perdita di qualità, in cui il segnale d'uscita del blocco 26 è portato direttamente a valle aggirando il blocco di de-quantizzazione 27.

Gli schemi rappresentati nelle figure 1 e 2 hanno carattere totalmente generale e generico e si applicano  
35 a segnali d'immagine digitalizzati di qualunque tipo.

L'operazione di quantizzazione esemplificata dal blocco 13 (le stesse considerazioni valgono naturalmente, in modo complementare, per il blocco di de-quantizzazione 27) prevedono che l'array di dati rappresentativi del segnale in ingresso - nel caso specifico, l'errore di predizione  $e(i)$  - sia suddiviso in blocchi di pixel. Quando il blocco di pixel contiene un solo pixel, la tecnica di quantizzazione che ne deriva è detta quantizzazione scalare; altrimenti è detta quantizzazione vettoriale o multidimensionale.

Per avere la minima complessità computazionale, si sceglie di preferenza di avere due pixel per blocco. In questo caso il quantizzatore è bidimensionale, ossia il vettore da quantizzare ha dimensione 2.

Quando il numero dei pixel lo rende possibile, il blocco è preferibilmente di forma quadrata. Tutto questo in modo da non privilegiare l'orientamento verticale rispetto a quello orizzontale, così da aumentare l'isotropia della compressione.

Nel caso di sequenze d'immagini, il blocco tridimensionale è preferibilmente di forma cubica per le ragioni già illustrate.

L'intensità dei pixel contenuti nel blocco costituisce una n-upla di coordinate in uno spazio n-dimensionale. Lo spazio n-dimensionale è partizionato in celle, ognuna contenente un punto di ricostruzione. Ad ogni punto di ricostruzione è assegnato un codice binario opportuno.

L'operazione di quantizzazione vettoriale o multidimensionale suscettibile di essere attuata nel blocco 13 consiste nel passare dalla rappresentazione binaria dell'intensità dei pixel contenuti nel blocco, al codice binario assegnato al punto di ricostruzione della cella selezionata in base ai pixel stessi.

Il quantizzatore vettoriale più semplice è quello

costituito da un quantizzatore scalare applicato n  
volte a ciascun pixel del blocco. Il quantizzatore  
vettoriale qui rappresentato consiste essenzialmente in  
un quantizzatore scalare applicato ad ogni pixel nel  
5 blocco con passo di quantizzazione calcolato in modo  
adattativo sulla base delle caratteristiche dei pixel  
stessi.

La tecnica di compressione è, come si è visto, di  
tipo lossy. Infatti il punto di ricostruzione è  
10 generalmente coordinato e diverso dal punto  
corrispondente alla n-upla di coordinate di partenza.  
La differenza è detta errore di quantizzazione.

Così come meglio illustrato nel seguito,  
progettando in modo adeguato il quantizzatore  
15 vettoriale, è possibile ottenere contemporaneamente la  
riduzione della ridondanza statistica e della  
ridondanza percettiva nei dati, pur mantenendo bassa la  
complessità computazionale.

In un quantizzatore vettoriale con celle uniformi  
20 e codice a lunghezza variabile per i punti di  
ricostruzione realizza la riduzione della ridondanza  
statistica dei dati. Questo avviene assegnando codici  
binari "corti" ai punti di ricostruzione appartenenti  
alle celle selezionate più frequentemente, in modo  
25 analogo a quanto si fa nella tecnica di compressione di  
Huffman.

Il quantizzatore vettoriale con celle non uniformi  
e codice a lunghezza fissa per i punti di ricostruzione  
ha asintoticamente le stesse prestazioni di riduzione  
30 della ridondanza statistica se costruito in modo  
opportuno. In particolare, le zone dello spazio  
multidimensionale selezionate più frequentemente sono  
partizionate con celle più piccole e tra loro più  
vicine.

35 La compressione è dovuta al fatto che la

rappresentazione binaria associata al punto di ricostruzione richiede meno bit della rappresentazione binaria degli elementi del vettore da quantizzare.

Per quanto riguarda il quantizzatore vettoriale  
 5 qui presentato, dopo misure sperimentali si osserva che per un spazio  $n$ -dimensionale le celle devono essere addensate lungo la retta descritta dalla equazione:

$$x_1 = x_2 = x_3 = \dots x_n,$$

in cui  $x_i$  è la coordinata  $i$ -esima, con  $i$  che va da  
 10 1 a  $n$ . Si tratta, in pratica, della diagonale  $n$ -dimensionale.

Le misure sperimentali sono giustificate dal fatto che i contorni sono una piccola parte dell'immagine. Nel resto dell'immagine, l'intensità dei pixel del  
 15 blocchetto è circa uniforme, il che significa che nello spazio multidimensionale la coordinata corrispondente è vicina alla retta indicata dall'equazione.

Si può inoltre osservare che, mentre il numero dei pixel dell'immagine cresce con  $N^2$ , vale a dire con il  
 20 quadrato della lunghezza del lato (supponendo che l'immagine sia circa quadrata, con lato di  $N$  pixel), il numero dei pixel dei contorni cresce solamente con  $N$ . Basandosi su quest'osservazione, si può prevedere che il quantizzatore vettoriale così progettato si dimostra  
 25 sempre più efficiente al crescere della risoluzione dell'immagine.

In altre parole, al crescere della risoluzione, percentualmente i contorni sono sempre più rari. Le  
 30 celle più grandi, lontane dalla diagonale  $n$ -dimensionale, con errore di quantizzazione più grande, sono selezionate percentualmente di meno. Si può concludere che il rapporto segnale/rumore di quantizzazione dell'immagine cresce.

Sulla base delle premesse sopra richiamate è  
 35 possibile, sempre con riferimento ad una quantizzazione



vettoriale non uniforme, parlare di riduzione della ridondanza percettiva.

Come già si è detto, il sistema visivo umano dimostra una ridotta sensibilità alle alte frequenze  
5 contenute nell'immagine, che possono quindi essere quantizzate in modo più grossolano o rozzo.

Per sfruttare in questo modo tale proprietà del sistema visivo, si passa al dominio delle frequenze tramite una trasformazione o trasformata (DCT o DWT) di  
10 un blocco di pixel. Operazione, questa, che di solito risulta piuttosto onerosa in termini di complessità computazionale.

Un modo più semplice per sfruttare questa proprietà è basato sul fatto di operare nel dominio  
15 spaziale piuttosto che nel dominio delle frequenze.

Tutto questo basandosi sull'idea di fondo di quantizzare più rozzamente le immagini vicino ai contorni, dove si trovano alte frequenze: si tratta in pratica di un effetto di mascheratura che si ha in  
20 corrispondenza dei contorni.

Quantizzare in modo più rozzo significa aumentare il passo di quantizzazione. In effetti, il passo di quantizzazione può essere tanto più grande quanto più netto è il contorno e la nettezza del contorno può  
25 essere misurata in diversi modi.

Ad esempio, se il blocco contiene solo due pixel (come in un'implementazione di minima complessità), la nettezza del contorno è semplicemente la differenza assoluta delle intensità luminose dei due pixel. Se il  
30 blocco contiene più di due pixel, una misura semplice e veloce può essere la differenza assoluta tra l'intensità più alta e l'intensità più bassa presenti nel blocco.

Dopo aver calcolato la nettezza del contorno  
35 (secondo criteri noti, sulla base delle premesse sopra

indicate) la si quantizza in modo da suddividere i contorni in più classi. Ad esempio nel caso di un blocco di due pixel:  $E = \text{abs}(p1 - p2)$ . Se  $p1$  e  $p2$  vanno da 0 a 255, la nettezza del contorno (edge)  $E$  va da 0 a 255. Quantizzando  $E$  con un passo  $E_{sq} = \max(E)/3$  si calcola  $E_Q = \text{round}(E/E_{sq})$  ottenendo quattro classi di contorni perché  $E_Q$  va da 0 a 3.

In alternativa, è possibile utilizzare una tabella di soglie per classificare  $E$  assegnando di conseguenza  $E_Q$ .

La classe 0 corrisponde al contorno meno netto (area piatta), la classe 3 al contorno più netto. Il passo di quantizzazione dei pixel del blocco è scelto tanto più alto quanto più netto è il contorno.

Il passo di quantizzazione  $Q$  può essere letto da una tabella indicizzata secondo la classe del contorno  $E_Q$ . In alternativa, il passo di quantizzazione  $Q$  può essere calcolato in base al parametro  $E_Q$ .

Ad esempio, si può porre semplicemente  $Q = m \cdot E_Q + q$  con  $M$  e  $q$  costanti, opportunamente scelte. In alternativa, ed in modo preferito, si può porre  $Q = m \cdot (t^{E_Q}) + q$ . Preferibilmente, per mantenere bassa la complessità computazionale,  $q = 0$  mentre  $t = 2$ .

Più in generale, e preferibilmente, la funzione che permette di calcolare  $Q$  da  $E_Q$  è definita in modo che il valore di  $Q$  sia moltiplicato per un valore intero se  $E_Q$  cresce e sia diviso per un valore intero se  $E_Q$  diminuisce. Questo permette di semplificare il successivo calcolo del codice binario (blocco 14 della figura 1) da fornire in uscita dal risultato della quantizzazione vettoriale (blocco 13 della stessa figura 1).

La quantizzazione vettoriale è quindi ridotta all'applicazione ripetuta di un quantizzatore scalare con passo di quantizzazione  $Q$  calcolato in modo

adattativo per sfruttare l'effetto di mascheratura del contorno. In particolare, l'intensità di ciascun pixel del blocco  $p_i$  è quantizzato nel seguente modo:  $p_{iQ} = \text{round}(p_i/Q)$ .

- 5        Il livello di ricostruzione (per invertire operazione di quantizzazione - blocco 27 della figura 2) è posto semplicemente pari a  $p_{iR} = p_{iQ} * Q$ .

10       L'insieme delle intensità quantizzate  $p_{iQ}$  del blocco è utilizzato per accedere ad una tabella nel quale è specificato il codice binario a lunghezza costante o variabile associato al livello di costruzione precedentemente individuato. Preferibilmente, tale codice è a lunghezza costante.

- 15       In alternativa, tale codice binario è ricavato con operazione aritmetiche binarie dalla rappresentazione binaria delle intensità quantizzate del blocco opportunamente raggruppate.

20       Si apprezzerà che il quantizzatore vettoriale che ne risulta ha celle più addensate lungo la diagonale n-dimensionale. Si realizza così contemporaneamente la riduzione della ridondanza statistica e della ridondanza percettiva.

- 25       Un'altra proprietà del sistema visivo, è la minore sensibilità all'aumentare dell'intensità luminosa. Questa proprietà è già ben sfruttata dalla relazione esponenziale che lega le componenti cromatiche lineari a quelle corrette non lineari, con l'uso del cosiddetto fattore gamma.

30       Le componenti corrette sono quelle normalmente usate. Si verifica sperimentalmente che tali componenti corrette presentano ancora in parte la menzionata ridondanza percettiva.

- 35       In particolare, si calcola il livello medio  $M$  di luminosità dei pixel nel blocco. Tale livello è quantizzato con passo  $M_{SQ}$  in modo da suddividere i

blocchi in più classi:  $M_Q = \text{round}(M/M_{SQ})$ . Ad esempio, se  $M_{SQ} = \max(M_Q)/2$ ,  $M_Q$  può valere 0, 1 oppure 2.

In alternativa, è possibile usare una tabella di soglia per classificare opportunamente  $M$  assegnando di conseguenza  $M_Q$ .

Il passo di quantizzazione  $Q$  può essere aumentato o diminuito di una quantità proporzionale a  $M_Q$  attraverso un'opportuna costante sfruttando così l'effetto di mascheramento dell'alta intensità luminosa.

In alternativa, è possibile usare una tabella per ricavare  $Q$  da  $M_Q$ . E' anche possibile calcolare  $Q$  direttamente da  $E_Q$  e  $M_Q$  presi insieme mediante un'apposita formula o tabella, quest'ultima di solito implementata sotto forma di una cosiddetta look up table (LUT).

Così come segnalato in precedenza, la funzione che permette di calcolare  $Q$  da  $E_Q$  è fatta di preferenza in modo tale che il valore di  $Q$  è moltiplicato per un valore intero se  $M_Q$  cresce e sia diviso per un valore intero se  $M_Q$  diminuisce.

Il quantizzatore vettoriale che ne risulta, ricavato dal precedente, ha celle più addensate lungo la diagonale  $n$ -dimensionale ( $x_1 = x_2 = x_3 = \dots = x_n$ ). In particolare, l'addensamento è maggiore all'inizio della diagonale ( $x_1 = x_2 = x_3 = \dots = x_n = c$  con  $c$  piccolo) e minore alla fine ( $x_1 = x_2 = x_3 = \dots = x_n = c$  con  $c$  grande).

Il quantizzatore vettoriale sin qui descritto è costituito da un quantizzatore scalare applicato a ciascun elemento del vettore da quantizzare. Il passo di quantizzazione è identico per tutti gli elementi del vettore ed è calcolato in base alle caratteristiche percettive del vettore stesso: nettezza del contorno, eventualmente presente e luminosità media.

I punti di ricostruzione di questo quantizzatore vettoriale sono disposti secondo un reticolo ortogonale: con celle quadrate, nel caso bidimensionale e con celle cubiche nel caso tridimensionale.

5 Per il caso bidimensionale è noto che il reticolo ottimale, con celle tutte uguali tra loro è quello con celle esagonali. Il motivo è che il massimo errore di quantizzazione è dovuto al punto che nella cella si trova più distante dal punto di ricostruzione. La cella  
10 ideale è quindi quella circolare e la cella esagonale è quella che meglio approssima la cella ideale pur ricoprendo interamente lo spazio da quantizzare.

Il quantizzatore con celle esagonali può essere realizzato da un quantizzatore con celle rettangolari  
15 in cui i punti di ricostruzione hanno coordinate che si incrementano con passo  $DX = 3/2 \cdot L$ ,  $DY = \sin(\pi/3) \cdot L$  con  $L$  prefissato e calcolato nel modo già descritto per il passo di quantizzazione. I punti di ricostruzione del reticolo rettangolare hanno coordinate  $X=n \cdot DX$ ,  $Y=m \cdot DY$ ,  
20 con  $n$  e  $M$  interi. I punti di ricostruzione del reticolo esagonale sono un sottoinsieme di questi e precisamente sono i punti per i quali  $o = n+m$  è pari (oppure è dispari).

In alternativa, ed in modo preferito, la cella è  
25 quadrata. In questo caso  $DX = DY = L$  con  $L$  prefissato è calcolato nel modo già descritto per il passo di quantizzazione. Se si prendono solo i punti di ricostruzione per i quali  $o = n+m$  è pari (oppure dispari) si ottiene un reticolo quadrato ruotato di  
30  $45^\circ$ , dunque un reticolo sostanzialmente assimilabile ad una configurazione a quinconce.

In alternativa è possibile usare un reticolo con celle esagonali ruotato di  $45^\circ$ . Questo in modo da allineare uno dei bordi delle celle alla diagonale  $n$ -  
35 dimensionale rispetto a cui è partizionato lo spazio da

quantizzare.

Il quantizzatore vettoriale fin qui illustrato è realizzato da un quantizzatore scalare applicato n volte (uno per ogni pixel del blocco dell'immagine) per  
 5 il quale il passo di quantizzazione  $Q$  è calcolato in modo adattativo in base alla nettezza  $E$  del contorno presente nel blocco ed in base alla sua intensità luminosa media  $M$ . Si realizza in questo modo contemporaneamente la riduzione della ridondanza  
 10 statistica e percettiva.

Un tale quantizzatore può essere applicato ad array di pixel corrispondenti alla luminanza, alla crominanza o ad un determinato colore (R, G o B).

Nel caso di componenti cromatiche sottocampionate  
 15 (come nel caso di segnali video nei formati YUV, YIQ o JDbDr in formato 4:2:2) e intercalati (come nel caso di segnale video digitale YCbCr ITU-R BT.601), il blocco dei pixel va opportunamente trattato, riordinando e demultiplexando le componenti cui applicare il  
 20 quantizzatore vettoriale.

Ad esempio, prendendo una coppia di pixel a 8 bit relativi alla stessa componente cromatica, questa è sostituita da un indice a 8 bit che individua uno dei  
 25 punti di ricostruzione che sono addensati intorno alla diagonali bidimensionale. Il fattore di compressione è in questo caso di 2:1. E' evidente, che utilizzando 8 bit, che sono  $2^8=256$  celle nello spazio bidimensionale corrispondente a ciascuna coppia di pixel. Un caso particolare è rappresentato da dati organizzati secondo  
 30 un cosiddetto Bayer Pattern (si veda al riguardo la figura 3, che si riferisce ad uno schema di scansione della componente verde del Bayer Pattern per un sistema di codifica/decodifica DPCM adattativo).

I dati in formato Bayer rappresentano ovviamente  
 35 un'approssimazione delle componenti cromatiche di una

scena acquisibili da un sensore digitale. La qualità dell'immagine finale è strettamente legata agli algoritmi di ricostruzione/interpolazione del colore. Nel momento in cui ci si pone l'obiettivo di attuare una funzione di compressione è comunque importante, nel caso di una compressione con perdite (lossy), cercare di preservare un'alta fedeltà rispetto ai dati originali. Piccole alterazioni potrebbero infatti alterare/peggiore la qualità dell'immagine RGB finale con effetti quali falsi colori, un cosiddetto aliasing diffuso, ecc. E' quindi importante utilizzare tecniche che tengano conto delle particolare struttura, sfruttando appunto la correlazione cromatica dei diversi canali.

Seguendo infatti un approccio globale di tipo tradizionale (ad esempio di tipo JPEG) le transizioni tra i pixel di colore differente sarebbero codificate come componenti ad alta frequenza e conseguentemente alterate in modo marcato. D'altra parte, se si pensa ad applicazioni di costo ridotto, la tecnica deve necessariamente sfruttare solo poche risorse di calcolo, la ricostruzione essendo preferibilmente prevista in remoto.

Nel caso di un array contenente un pattern di Bayer, il blocco di dimensioni  $4 \times 2$  contiene il seguente schema di componenti cromatiche:

riga 1 =  $G_1 R_1 G_2 R_2$

riga 2 =  $B_1 G_3 B_2 G_4$

utilizzando il quantizzatore vettoriale bidimensionale le coppie cui lo stesso può essere applicato sono  $\langle R_1 R_2 \rangle$ ,  $\langle B_1 B_2 \rangle$  insieme a  $\langle G_1 G_2 \rangle$   $\langle G_3 G_4 \rangle$  oppure  $\langle G_1 G_3 \rangle$ ,  $\langle G_2 G_4 \rangle$ .

Sperimentalmente si osserva che, quantizzando vettorialmente  $\langle G_1, G_2 \rangle$ ,  $\langle G_3, G_4 \rangle$ , la qualità migliora. Infatti, nell'altro caso i pixel che fanno parte della

coppia sono più distanti ed è meno probabile che abbiano intensità simili, cioè che il punto di coordinate corrispondenti sia presso la diagonale dove la quantizzazione è più fine e l'errore è minore.

5        Nel caso di immagini in formato RGB, è di solito conveniente operare un cambiamento di coordinate allo spazio colore YCbCr (oppure ad uno degli analoghi spazi in cui l'informazione di luminanza è separata da quella di cromaticità, ossia YUV, YIQ o YDbDr).

10       Per una generale rassegna sulle caratteristiche di tali spazi cromatici, nonché degli altri spazi cromatici cui si fa riferimento nell'ambito della presente descrizione, si possono utilmente consultare i documenti:

- 15       - R.C. Gonzales, R.E. Woods, *Digital Image Processing*, Addison Wesley, 1993;  
          - W.B. Pennebaker, J.L. Mitchell, *Jpeg, still image data compression standard*, Van Nostrand Reinhold, 1992; e  
          - D. Taubman, M. Marcellin *JPEG2000 Image Compression Fundamentals*, The Kluwer Int. Series in Eng. and  
 20       Computer Science Volume 642 Hardbound, ISBN 0-7923-7519-X, Nov. 2001.)

         I piani di cromaticità sono quindi sottocampionati orizzontalmente (formato 4:2:2) oppure sia  
 25       orizzontalmente sia verticalmente (formato 4:2:0). Eventualmente, si può far precedere l'operazione di sottocampionamento da un filtraggio passa-basso per ridurre gli effetti di aliasing, soprattutto nel caso di materiale video intercalato e non progressivo.

30       Il piano di luminanza è quindi compresso applicando il quantizzatore vettoriale bidimensionale alle coppie adiacenti (orizzontalmente o verticalmente) di pixel. Ad esempio: dalla riga =  $Y_1Y_2Y_3Y_4$  si passa alle coppie  $\langle Y_1Y_2 \rangle$ ,  $\langle Y_3Y_4 \rangle$  poi quantizzate  
 35       vettorialmente. I piani di cromaticità sono elaborati



in modo analogo ma separatamente.

In alternativa, ed in modo preferito, il sottocampionamento dei piani di cromaticità avviene secondo una configurazione a quinconce (cioè a scacchiera), operando in modo che i due piani di cromaticità si possano intercalare perfettamente: riga 1 =  $U_1V_1U_2V_2$ , riga 2 =  $V_3U_3V_4U_4$ . La quantizzazione vettoriale va poi applicata alle coppie  $\langle U_1U_3 \rangle$ ,  $\langle U_2U_4 \rangle$  e  $\langle V_1V_3 \rangle$ ,  $\langle V_2V_4 \rangle$ .

In alternativa ma con qualità leggermente inferiori, si possono utilizzare le seguenti coppie:  $\langle U_1U_2 \rangle$ ,  $\langle U_3U_4 \rangle$  e  $\langle V_1V_2 \rangle$ ,  $\langle V_3V_4 \rangle$ . La qualità inferiore è dovuta alla maggiore distanza spaziale fra i pixel delle coppie, che rende più probabile una minore correlazione. La coppia da quantizzare vettorialmente si colloca quindi lontano dalla diagonale multidimensionale ed è quantizzata con un più consistente errore di quantizzazione.

Questo reticolo di sottocampionamento a scacchiera risulta maggiormente isotropo rispetto al caso 4:2:2, non privilegiando i bordi orizzontali. Inoltre, il perfetto intercalamento fra le componenti fa sì che ad ogni pixel di luminanza sia associata sempre una componente di cromaticità, invece di avere alternativamente pixel per i quali è specificata la terna YUV e pixel per i quali sia solo il valore Y. Questo permette di ridurre gli artefatti dovuti alla successiva fase di interpolazione e ricostruzione dell'immagine RGB originaria.

Così come si è visto, la quantizzazione vettoriale di componenti cromatiche intercalate avviene raggruppando ciascuna componente nel blocco in un vettore di opportuna dimensione e quantizzandola separatamente. Preferibilmente il vettore ha dimensione minima pari a 2.

Volendo decomprimere l'immagine, si deve sostituire al codice binario nel punto di ricostruzione (fornito in uscita dal blocco 26 della figura 2) le coordinate del punto di ricostruzione stesso. Nel caso  
 5 di componente cromatiche sottocampionate ed intercalate (multiplexate) è poi necessario separare le componenti (demultiplexing) ed interpolare.

Si osserva sperimentalmente che l'errore di quantizzazione vettoriale dà origine a pattern colorati  
 10 visibili. Ciò avviene soprattutto nelle aree uniformi dell'immagine. La causa è da individuarsi principalmente nel metodo di interpolazione che necessariamente fa uso di pixel adiacenti della stessa componente cromatica ma affetti da errore di  
 15 quantizzazione ripetuto e regolare.

Una possibile soluzione del problema consiste nel cercare di introdurre del cosiddetto dithering per spezzare la regolarità dell'errore di quantizzazione.

In pratica, si aggiunge intenzionalmente un  
 20 (basso) livello di rumore in modo da evitare il pattern colorato. Lo svantaggio è che la qualità dell'immagine diminuisce leggermente in quanto l'immagine appare leggermente granulosa.

Una seconda possibile soluzione consiste  
 25 nell'applicare tecniche di sagomatura del rumore (noise shaping), tenendo conto - durante la quantizzazione - dell'errore di quantizzazione precedente. Lo svantaggio di questa soluzione è che si aumenta (anche se di poco) la complessità del quantizzatore.

30 Una soluzione particolarmente vantaggiosa consiste nel prevedere per le aree uniformi (con  $E=0$ ) la quantizzazione più fine possibile con il minimo passo di quantizzazione ( $Q=1$ ).

Ad esempio, sviluppando l'esempio visto in  
 35 precedenza (passaggio da RGB a YCbCr modificato), per

ciascuna coppia  $\langle c_1, c_2 \rangle$ , se  $c_1 = c_2 = c$  (8 bit), allora si manda la sequenza di bit "1" + 8 bit di  $c$ , altrimenti, se  $c_1$  è diverso da  $c_2$ , si manda la sequenza di bit "0" più 8 bit corrispondenti all'indice  $VQ$  ( $c_1 c_2$ ). In totale si passa dai 16 bit della coppia ( $c_1$ ,  $c_2$ ) a 8 bit.

Questa soluzione si colloca naturalmente nello schema delineato di quantizzatore vettoriale. Infatti, impostare al minimo il passo di quantizzazione ( $Q=1$ ) per le aree uniformi (dove  $E=0$ ) significa che i punti di ricostruzione del quantizzatore vettoriale, oltre ad essere addensati vicino alla diagonale  $n$ -dimensionale, si collocano anche sulla diagonale stessa. Questo è importante per evitare pattern colorati visibili.

Tali pattern fanno peraltro parte di quella categoria di artefatti per la quale le misure standard di qualità (come il fattore PSNR) non corrispondono alla qualità percepita. Infatti, le misure standard di qualità sono basate sull'intensità dell'artefatto e senza tener conto della sua influenza sul sistema visivo. Nel presente caso i pattern in questione hanno una intensità piccola, ma essendo regolari sono suscettibili di risultare ben visibili ed essere quindi percepiti dall'utilizzatore.

L'operazione di quantizzazione vettoriale o multidimensionale qui descritta consiste nel passare dalla rappresentazione binaria dell'intensità dei pixel contenuti nel blocco al codice binario assegnato al punto di ricostruzione della cella selezionata in base ai pixel stessi.

Cambiando punto di vista si può affermare, con sostanziale aderenza alla realtà che il blocco dei pixel è codificato come una sorta di unico "superpixel" avente intensità specificata dal codice binario associato al punto di ricostruzione selezionato.

Sperimentalmente, si evidenzia una residua ridondanza statistica dei dati, suscettibile di essere ulteriormente ridotta in modo lossy applicando nuovamente la soluzione qui presentata oppure in modo  
 5 senza perdite (lossless) concatenando uno dei metodi di compressione entropica già presentati.

L'efficienza della compressione lossy iterata dipende essenzialmente dal modo con cui i codici binari sono assegnati a ciascun punto di ricostruzione. In  
 10 particolare, prendendo due punti di ricostruzione, l'assegnazione deve essere fatta in modo che ad una distanza minore nello spazio n-dimensionale corrisponda una differenza minore dei valori binari corrispondenti. Supponendo che una tale assegnazione sia stata fatta,  
 15 si comprende che il valore di un superpixel ha un valore simile a quello dei superpixel adiacenti proprio a causa della residua ridondanza statistica. I superpixel possono quindi essere posti in vettori multidimensionali e quantizzati con un procedimento  
 20 simile a quello già illustrato.

Prendendo riferimento nuovamente all'esempio del Bayer pattern, nel caso di componenti cromatiche intercalate si passa da un Bayer pattern ad un "super" Bayer pattern. Infatti, ogni blocco 4x2 del blocco  
 25 Bayer di partenza (riga 1 =  $G_1R_1G_2R_3$ , riga 2 =  $B_1G_3B_2G_4$ ) è codificato in un blocco di superpixel 2x2 (riga 1 =  $G'R^*$ , riga 2 =  $B*G''$ ), mediante la quantizzazione vettoriale bidimensionale ( $R^*=VQ\langle R_1, R_2 \rangle$ ,  $B^*=VQ\langle B_1, B_2 \rangle$ ,  $G'=BQ\langle G_1, G_2 \rangle$ ,  $G''=VQ\langle G_3, G_4 \rangle$ ).

30 E' evidente la possibilità di iterare il metodo sul super Bayer pattern così ottenuto. Dal punto di vista della quantizzazione vettoriale, questo significa aumentare la dimensione del quantizzatore vettoriale, dato che si agisce su blocchi che fanno riferimento a  
 35 porzioni sempre più grandi dell'array di dati di

partenza.

E' altresì evidente la possibilità di iterare la compressione anche nel caso relativo ad YCbCr modificato.

5        Per conseguire un'ulteriore riduzione della ridondanza statistica residua è possibile ricorrere ad una codifica entropica.

A titolo esemplificativo, il metodo più semplice consiste nell'individuare superpixel adiacenti con lo  
10        stesso valore. Tale sequenza di superpixel è allora ridotta ad un unico esemplare del superpixel, preceduto o seguito da un contatore che ne indica il numero di ripetizioni. Si tratta in altre parole di una applicazione della tecnica nota come codifica Run-  
15        Length Encoding o RLE.

L'applicazione di metodi più evoluti è più efficace con un'opportuna assegnazione dei codici binari e i punti di ricostruzione (blocco 26 della  
20        figura 2). In particolare, l'assegnazione ottimale dei codici segue la regola già illustrata in precedenza.

Prendendo due punti di ricostruzione, l'assegnazione deve essere fatta in modo che ad una distanza minore nello spazio n-dimensionale corrisponda una differenza minore dei valori binari corrispondenti.

25        Supponendo che sia stata fatta una tale assegnazione, si intuisce che il valore di un superpixel è predicibile in base al valore dei superpixel adiacenti. Nel caso più semplice, si predice il valore del superpixel in base al valore di quello  
30        che precede. Si codifica quindi l'errore di predizione con una tecnica sostanzialmente identificabile con una tecnica DPCM cui fanno riferimento appunto i diagrammi delle figure 1 e 2.

Lontano dai contorni (e quindi probabilmente per  
35        la maggior parte dei superpixel), l'errore di

predizione è piccolo. Valori piccoli di questo errore sono quindi codificabili con codici binari brevi, mentre valori grandi avranno codici binari lunghi, analogamente a quanto visto a proposito della  
 5 compressione di Huffman.

Ovviamente, nel caso di componenti cromatiche intercalate, la predizione più semplice di un determinato superpixel è fatta in base al valore del superpixel più vicino e relativo alla stessa componente  
 10 cromatica.

Nei casi più complessi, invece, la predizione di una componente cromatica può essere fatta in base a superpixel adiacenti appartenenti ad un'altra componente cromatica (così come è fatto normalmente  
 15 nei metodi di interpolazione cromatica).

Ad esempio, nel caso particolare in cui i dati codificati siano in formato Bayer, una leggera modifica nello schema di predizione di tipo DPCM permette di migliorare le prestazioni nel caso della codifica della  
 20 componente di verde.

In un Bayer pattern, infatti i pixel verdi sono presenti su ogni riga, mentre quelli blu e rossi sono distribuiti su righe alternate. Pertanto, la maggior vicinanza spaziale tra pixel verdi continui  
 25 appartenenti a righe successive, rispetto a pixel adiacenti alla stessa riga, si traduce in una maggior correlazione che, a sua volta, porta a errori di predizione minori. Questo almeno nel caso di zone di immagini in cui non sono presenti lati netti.

Uno schema di predizione che segua un ordine "a zig-zag", nel calcolo degli errori, permette di ottenere una compressione leggermente migliore rispetto allo schema classico, sia nel caso di compressione  
 30 lossless che in quello di compressione lossy.

L'entità di tale miglioramento (cui fa specifico  
 35

riferimento la fig. 3) dipende dalle caratteristiche  
 dell'immagine e aumenta al crescere della risoluzione.  
 La tabella 1 qui sotto riprodotta mostra i risultati  
 medi ottenuti su basi dati di immagini in Bayer pattern  
 5 di risoluzioni differenti e compresse usando sia  
 l'approccio di tipo DPCM adattativo, appena descritto  
 (I-DPCM), sia quello classico (DPCM standard o STD).

risolu- zione	algoritmo	Prestazioni compressione bit rate (bpp)						
		lossless	q=2	q=4	q=8	q=16	q=24	q=32
352x288	DPCM std	4,79	3,83	2,96	2,19	1,65	1,45	1,33
	I-DPCM	4,67	3,73	2,85	2,10	1,58	1,40	1,29
640x480	DPCM std	4,96	4,04	3,14	2,39	1,81	1,57	1,43
	I-DPCM	4,87	3,95	3,06	2,31	1,71	1,54	1,40
1000x800	DPCM std	3,57	2,74	2,07	1,62	1,34	1,24	1,18
	I-DPCM	3,44	2,62	1,98	1,55	1,30	1,21	1,15

10 Nella tabella, il valore q indica il passo  
 utilizzato nel caso di compressione lossy ottenuta  
 tramite quantizzazione uniforme.

Concatenando il sistema di codifica DPCM lossless  
 all'approccio basato sulla quantizzazione vettoriale, è  
 15 possibile ottenere una compressione migliore senza  
 perdite di qualità nell'uscita.

Fermo restando il fatto che il quantizzatore  
 vettoriale opera in modo soddisfacente con qualunque  
 tipo d'immagine dal punto di vista percettivo  
 20 (valutazione di qualità soggettiva), è possibile  
 migliorarne le prestazioni oggettive (ad esempio in  
 termini di rapporto segnale/rumore di picco o PSNR) nel  
 caso d'immagini contenenti contorni netti: un tipico  
 esempio è costituito da immagini create artificialmente  
 25 al computer, cartoni animati, pagine di testo acquisite  
 con scanner o no.

Per ottenere questo miglioramento si modifica la funzione che calcola il passo di quantizzazione  $Q$  in funzione della nettezza del contorno  $E$  e del livello medio di luminosità  $M$ . In particolare,  $Q$  è scelto  
 5 piccolo quando  $E$  è massimo (contorno netto). In definitiva, non si fa semplicemente crescere  $Q$  con  $E$  e  $M$  come visto prima ma  $Q$  raggiunge un suo massimo in corrispondenza di un valore intermedio di  $E$ .

Questo significa che i punti di ricostruzione del  
 10 reticolo di quantizzazione sono posti negli angoli lontani dalla diagonale multidimensionale. Questi angoli appartengono al cubo multidimensionale in cui a collocarsi il vettore corrispondente alla  $n$ -upla di coordinate corrispondenti ai pixel nel blocco.

15 In alternativa (ed in aggiunta alla strategia appena illustrata per il calcolo di  $Q$ ), è possibile fare in modo che  $Q$  sia piccolo quando uno dei pixel nel blocco ha il valore minimo o massimo ammissibile. Questo significa che i punti di ricostruzione del  
 20 reticolo di quantizzazione sono posti non solo negli angoli lontani dalla diagonale multidimensionale, ma anche lungo i lati del cubo multidimensionale in cui va a collocarsi il vettore corrispondente alla  $n$ -upla di coordinate corrispondenti ai pixel nel blocco.

25 I grafici delle figure 4a a 4f riproducono esempi di reticoli di punti di ricostruzione e celle di quantizzazione per quantizzatori vettoriali bidimensionali. In particolare, i grafici in questione (le cui scale tanto da ascissa quanto di ordinata sono  
 30 di per sé irrilevanti) si riferiscono rispettivamente ad un reticolo quadrato (figure 4a, 4b), ad un reticolo esagonale (figura 4c).

Le figure 4d, 4e e 4f si riferiscono invece a valori di errore ricavati riscontrati rispettivamente  
 35 con un quantizzatore  $Q$  scalare, un quantizzatore  $Q$



vettoriale e un quantizzatore  $Q$  vettoriale ottimale.

Le immagini delle figure da 5a a 5d stanno a dimostrazione delle prestazioni di un generico quantizzatore vettoriale bidimensionale applicato a  
 5 blocchi  $2 \times 1$  di un array contenente luminanza (fattore di compressione 50%, da 8 bpp a 4 bpp) rispetto ad un quantizzatore scalare.

In particolare, la figura 5a) riproduce un'immagine a 8 bit/pixel mentre la figura 4b) fa  
 10 vedere la statistica  $2Q$  per un blocco  $2 \times 1$ .

La figura 5c illustra la distribuzione delle celle dei punti di ricostruzione, mentre la figura 5d illustra l'andamento dell'errore di quantizzazione.

Le figure 5e e 5f) riproducono le celle e i punti  
 15 di ricostruzione nonché l'errore di quantizzazione con riferimento alle immagini riprodotte nelle figure 5g e 5h) inerenti a immagini su quattro bit per pixel riprodotte dopo quantizzazione con quantizzatore  $2Q$  uniforme e quantizzatore  $2Q$  non uniforme.

Le figure 6a a 6l stanno a dimostrazione delle  
 20 prestazioni di un quantizzatore vettoriale bidimensionale costruito secondo le modalità descritte in precedenza ed applicato a blocchi  $4 \times 2$  di un array contenente componenti cromatiche intercalate secondo il  
 25 Bayer pattern (fattore di compressione 56% da 8 bpp a 4,5 bpp) rispetto a un quantizzatore scalare e rispetto ad un quantizzatore vettoriale ma senza punti lungo la diagonale multi-dimensionale.

In particolare, si suppone di partire da  
 30 un'immagine RGB a 24 bit/per pixel (figura 6a) cui corrispondente in Bayer pattern su 8 bit/pixel riprodotto nella figura 5b.

La figura 5c illustra l'immagine RGB ricostruita dal Bayer pattern e la figura 6d riproduce in scala  
 35 ingrandita una parte della stessa immagine.

Le figure 6e e 6f illustrano, a titolo di confronto, l'immagine RGB da Bayer pattern compresso con quantizzatore scalare, mentre le figure 6g e 6h si riferiscono all'immagine RGB ricostruita da Bayer pattern compresso con quantizzatore vettoriale.

Le figure 6i e 6l si riferiscono poi ad un'immagine RGB ricostruita da Bayer pattern compresso con quantizzatore vettoriale ottimale.

La figura 7a evidenzia la forte correlazione fra le componenti  $G_1$  (ascissa) e  $G_2$  (ordinata) del Bayer pattern. Si vede quindi che la coppia  $\langle G_1, G_2 \rangle$  si colloca lungo la diagonale bidimensionale dello spazio di quantizzazione ed è quantizzata con un piccolo errore di quantizzazione.

Al contrario è evidente che questo non accade per le coppie  $\langle G, R \rangle$  - figura 7b -, per la coppia  $\langle G, B \rangle$  - figura 7c - e per la coppia  $\langle R, B \rangle$  - figura 7d. Per tali coppie il risultato della quantizzazione vettoriale percettiva sarebbe quindi affetto da un errore di quantizzazione maggiore.

Si sottolinea che anche in relazione alle figure 7a - 7d, così come per le figure 4a-4f, 5a-5h e 6a-6l per l'esatta definizione delle scale non è di per sé rilevante.

Naturalmente, fermo restando il principio dell'invenzione, i particolari di realizzazione e le forme di attuazione potranno essere ampiamente variati rispetto a quanto descritto ed illustrato, senza per questo uscire dall'ambito della presente invenzione, così come definita dalle rivendicazioni annesse.

### RIVENDICAZIONI

1. - Procedimento per convertire segnali video digitali organizzati in blocchi di pixel tra un primo formato (10, 30) ed un secondo formato (15, 25), detto  
5 secondo formato essendo un formato compresso tramite quantizzazione vettoriale (13, 27), caratterizzato dal fatto che detta quantizzazione vettoriale risulta dall'applicazione ripetuta di un quantizzatore scalare applicato ai pixel di detti blocchi con passo di  
10 quantizzazione (Q) determinato in modo adattativo in funzione delle caratteristiche dei pixel.

2. - Procedimento secondo la rivendicazione 1, caratterizzato dal fatto che detto passo di quantizzazione (Q) è determinato in modo da risultare  
15 crescente con la disuniformità di pixel nel blocco.

3. - Procedimento secondo la rivendicazione 2, caratterizzato dal fatto che detto passo di quantizzazione (Q) è determinato da una legge crescente secondo multipli.

20 4. - Procedimento secondo una qualsiasi delle rivendicazioni 1 a 3, caratterizzato dal fatto che comprende l'operazione di rilevare il livello di luminosità dei pixel nel blocco e l'operazione di determinare detto passo di quantizzazione (Q) in modo  
25 che tale passo è crescente in funzione di detto livello di luminosità.

5. - Procedimento secondo la rivendicazione 1, caratterizzato dal fatto che comprende le operazioni di:

- 30 - rilevare la disuniformità dei pixel nel blocco,  
- rilevare il livello di luminosità dei pixel nel blocco, e  
- determinare detto passo di quantizzazione (Q) in modo tale che detto passo di quantizzazione è prima  
35 crescente e poi decrescente in funzione di detta

disuniformità e di detto livello di luminosità.

6. - Procedimento secondo la rivendicazione 5, caratterizzato dal fatto che detto passo di quantizzazione (Q) è fatto crescere e decrescere  
5 secondo multipli e sottomultipli.

7. - Procedimento secondo una qualsiasi delle rivendicazioni 4 a 6, caratterizzato dal fatto che detta operazione di rilevare il livello di luminosità dei pixel nel blocco è condotta rilevando il livello  
10 medio di luminosità dei pixel nel blocco.

8. - Procedimento secondo una qualsiasi delle precedenti rivendicazioni, caratterizzato dal fatto che nel passaggio da detto primo formato (10) a detto secondo formato (15) detti segnali video digitali sono  
15 sottoposti ad almeno un'operazione scelta nel gruppo costituito da:

- sottocampionamento,
- filtraggio passa-basso in funzione anti-aliasing prima del sottocampionamento, e  
20 - multiplexing di almeno una parte dei dati digitali necessari per la rappresentazione dell'immagine.

9. - Procedimento secondo una qualsiasi delle precedenti rivendicazioni, applicato a segnali video  
25 digitali comprendenti componenti cromatiche intercalate, caratterizzato dal fatto che nel passaggio da detto primo formato (10) a detto secondo formato (15) comprende l'operazione di riordinare i pixel in ciascun blocco da quantizzare componendoli in un  
30 vettore tale per cui le componenti cromatiche intercalate sono quantizzate separatamente.

10. - Procedimento secondo la rivendicazione 1, caratterizzato dal fatto che comprende l'operazione di identificare, nell'ambito di detti segnali video  
35 digitali, blocchi di pixel uniformi ( $E = 0$ ) scegliendo

per tali blocchi di pixel uniformi un passo di quantizzazione minimo ( $Q = 1$ ), nell'ambito dei passi di quantizzazione adottati in detta quantizzazione vettoriale.

5        11. - Procedimento secondo una qualsiasi delle precedenti rivendicazioni, caratterizzato dal fatto che detti segnali video digitali in detto secondo formato sono espressi sotto forma di codici binari associati a rispettivi segnali quantizzati (13, 27) e dal fatto che  
10        comprende l'operazione di svolgere una funzione di predizione (29) di detti codici binari.

12. - Procedimento secondo la rivendicazione 11, caratterizzato dal fatto che detta funzione di predizione dei codici binari è realizzata secondo uno  
15        schema DPCM.

13. - Procedimento secondo una qualsiasi delle precedenti rivendicazioni, caratterizzato dal fatto che nel passaggio da detto primo formato (1) a detto secondo formato (15) il segnale compresso tramite  
20        quantizzazione vettoriale è sottoposto a codifica entropica.

14. - Procedimento secondo la rivendicazione 13, caratterizzato dal fatto che detta codifica entropica è condotta con una tecnica scelta nel gruppo costituito  
25        da: Run Length Encoding, Huffman, aritmetica.

15. - Procedimento secondo una qualsiasi delle precedenti rivendicazioni, caratterizzato dal fatto che detta quantizzazione vettoriale è una quantizzazione vettoriale multidimensionale risultante dalla  
30        concatenazione di una pluralità di quantizzazioni vettoriali ciascuna risultante dall'applicazione ripetuta di una quantizzazione scalare.

16. - Procedimento secondo la rivendicazione 15, caratterizzato dal fatto che per ognuno di detti  
35        quantizzatori vettoriali concatenati, i codici binari

dei punti di ricostruzione sono assegnati in modo che punti di ricostruzione con distanza piccola nello spazio multidimensionale hanno codici binari con differenza ridotta.

5        17. - Procedimento secondo la rivendicazione 1, caratterizzato dal fatto che detto quantizzatore vettoriale è realizzato con l'applicazione di due quantizzatori scalari con passo di quantizzazione scalato con costanti  $2/3$  e  $\sin(\Pi/3)$ , rispettivamente.

10       18. - Procedimento secondo la rivendicazione 17, caratterizzato dal fatto di definire i punti di ricostruzione permessi in modo che questi formino un reticolo esagonale.

15       19. - Procedimento secondo la rivendicazione 1, caratterizzato dal fatto che comprende le operazioni di:

- identificare un valore di nettezza del contorno in ciascuno di detti blocchi di pixel, suddividendo i contorni in più classi,

20       - attribuire valori di detto passo di quantizzazione (Q) differenziati in funzione delle classi così determinate.

25       20. - Procedimento secondo la rivendicazione 19 caratterizzato dal fatto che detto passo di quantizzazione (Q) è determinato secondo almeno una legge scelta nel gruppo costituito da:

-  $m \cdot E_Q + Q$  e

-  $Q = m \cdot (t \wedge E_Q) + q$ , dove

$m$  e  $q$  sono costanti, selettivamente determinate, e

30        $E_Q$  è un indice che identifica detta classe di contorno.

21. - Procedimento secondo una qualsiasi delle rivendicazioni 4 a 7, caratterizzato dal fatto che comprende le operazioni di:

35       - suddividere detto livello di luminosità dei

pixel nel blocco in più classi, e

- selezionare il valore di detto passo di quantizzazione (Q) in modo differenziato in funzione delle classi così determinate.

5        22. - Procedimento secondo una qualsiasi delle precedenti rivendicazioni, caratterizzato dal fatto che detti segnali video in detto primo formato (10) sono segnali generati secondo un cosiddetto Bayer pattern, ordinabili in blocchi di dimensioni 4x2 aventi lo  
10 schema di componenti cromatiche:

riga 1 =  $G_1 R_1 G_2 R_2$ ,

riga 2 =  $B_1 G_3 B_2 G_4$ ,

e dal fatto che detta quantizzazione vettoriale è una quantizzazione vettoriale bidimensionale applicata alle  
15 coppie  $\langle R_1, R_2 \rangle$ ,  $\langle B_1, B_2 \rangle$ , insieme a  $\langle G_1, G_2 \rangle$ ,  $\langle G_3, G_4 \rangle$  oppure  $\langle G_1, G_3 \rangle$ ,  $\langle G_2, G_4 \rangle$ .

23.- Procedimento secondo una qualsiasi delle rivendicazioni 1 a 21, caratterizzato dal fatto che detti segnali video digitali in detto primo formato  
20 sono segnali digitali video in formato RGB e dal fatto che i segnali video digitali in detto secondo formato sono soggetti ad un cambiamento di coordinate ad uno spazio colore scelto dal gruppo costituito da YCbCr, YUV, UIQ o YDbDr.

25        24. - Procedimento secondo la rivendicazione 1 o la rivendicazione 23, caratterizzato dal fatto che detto quantizzatore vettoriale è applicato alle coppie adiacenti di pixel del piano di luminanza.

25. - Procedimento secondo la rivendicazione 8,  
30 caratterizzato dal fatto che, in detti segnali video digitali in detto secondo formato, i piani di cromaticanza sono sottocampionati secondo un pattern a quinconce.

26. - Sistema per convertire segnali video  
35 digitali organizzati in blocchi di pixel tra un primo

formato (10, 30) ed un secondo formato (15, 25), detto secondo formato essendo un formato compresso tramite quantizzazione vettoriale (13, 27), il sistema comprendendo almeno un convertitore scelto fra un  
 5 codificatore (13, 14) ed un decodificatore (26, 27) ed essendo caratterizzato dal fatto che detto convertitore è configurato per una quantizzazione vettoriale risultante dall'applicazione ripetuta di un  
 10 quantizzatore scalare applicato ai pixel di detti blocchi con passo di quantizzazione (Q) determinato in modo adattativo in funzione delle caratteristiche dei pixel.

27. - Sistema secondo la rivendicazione 26, caratterizzato dal fatto che detto convertitore è  
 15 configurato per determinare detto passo di quantizzazione (Q) in modo che tale passo di quantizzazione risulti crescente con la disuniformità di pixel nel blocco.

28. - Sistema secondo la rivendicazione 27, caratterizzato dal fatto che detto convertitore è  
 20 configurato per determinare detto passo di quantizzazione (Q) con una legge crescente secondo multipli.

29. - Sistema secondo una qualsiasi delle rivendicazioni 26 a 28, caratterizzato dal fatto che  
 25 detto convertitore è configurato per rilevare il livello di luminosità dei pixel nel blocco e determinare detto passo di quantizzazione (Q) in modo che tale passo è crescente in funzione di detto livello  
 30 di luminosità.

30. - Sistema secondo la rivendicazione 26, caratterizzato dal fatto che detto convertitore è configurato per:

- rilevare la disuniformità dei pixel nel blocco,
- 35 - rilevare il livello di luminosità dei pixel nel



blocco, e

- determinare detto passo di quantizzazione (Q) in modo tale che detto passo di quantizzazione è prima crescente e poi decrescente in funzione di detta  
5 disuniformità e di detto livello di luminosità.

31. - Sistema secondo la rivendicazione 30, caratterizzato dal fatto che detto convertitore è configurato per far crescere e decrescere detto passo di quantizzazione (Q) secondo multipli e sottomultipli.

10 32. - Sistema secondo una qualsiasi delle rivendicazioni 29 a 31, caratterizzato dal fatto che detto convertitore è configurato per rilevare il livello di luminosità dei pixel nel blocco rilevando il livello medio di luminosità dei pixel nel blocco.

15 33. - Sistema secondo una qualsiasi delle precedenti rivendicazioni 26 a 32, caratterizzato dal fatto che detto convertitore è un codificatore (11 a 14) configurato per sottoporre detti segnali video digitali convertiti da detto primo formato (10) e detto  
20 secondo formato (15) ad almeno un'operazione scelta nel gruppo costituito da:

- sottocampionamento,
- filtraggio passa-basso in funzione anti-aliasing prima del sottocampionamento, e
- 25 - multiplexing di almeno una parte dei dati digitali necessari per la rappresentazione dell'immagine.

34. - Sistema secondo una qualsiasi delle precedenti rivendicazioni 26 a 33, applicato a segnali  
30 video digitali comprendenti componenti cromatiche intercalate, caratterizzato dal fatto che detto codificatore (11 a 14) è configurato per riordinare i pixel in ciascun blocco da quantizzare componendoli in un vettore tale per cui le componenti cromatiche  
35 intercalate sono quantizzate separatamente.

35. - Sistema secondo la rivendicazione 26, caratterizzato dal fatto che detto codificatore (11 a 14) è configurato per identificare, nell'ambito di detti segnali video digitali, blocchi di pixel uniformi (E = 0) scegliendo per tali blocchi di pixel uniformi un passo di quantizzazione minimo ( $Q = 1$ ), nell'ambito dei passi di quantizzazione adottati in detta quantizzazione vettoriale.

36. - Sistema secondo una qualsiasi delle precedenti rivendicazioni 26 a 35, caratterizzato dal fatto che detto convertitore è configurato in modo tal per cui detti segnali video digitali in detto secondo formato sono espressi sotto forma di codici binari associati a rispettivi segnali quantizzati (13, 27) e dal fatto che lo stesso convertitore è configurato (11, 29) per svolgere una funzione di predizione (29) di detti codici binari.

37. - Sistema secondo la rivendicazione 36, caratterizzato dal fatto che detta funzione di predizione dei codici binari è realizzata secondo uno schema DPCM.

38. - Sistema secondo una qualsiasi delle precedenti rivendicazioni 26 a 37, caratterizzato dal fatto che detto convertitore è (11 a 14; 26 a 29) è configurato per sottoporre i segnali convertiti da detto primo formato (1) e detto secondo formato (15) ad una funzione di codifica o decodifica entropica.

39. - Sistema secondo la rivendicazione 38, caratterizzato dal fatto che detta codifica entropica è condotta con una tecnica scelta nel gruppo costituito da: Run Length Encoding, Huffman, aritmetica.

40. - Sistema secondo una qualsiasi delle precedenti rivendicazioni 26 a 39, caratterizzato dal fatto che detto convertitore è configurato per una quantizzazione vettoriale costituita da una

quantizzazione vettoriale multidimensionale risultante dalla concatenazione di una pluralità di quantizzazioni vettoriali ciascuna risultante dall'applicazione ripetuta di una quantizzazione scalare.

5        41. - Sistema secondo la rivendicazione 40, caratterizzato dal fatto che detto convertitore è configurato per assegnare i codici binari dei punti di ricostruzione per ognuno di detti quantizzatori vettoriali concatenati in modo che punti di  
10 ricostruzione con distanza piccola nello spazio multidimensionale hanno codici binari con differenza ridotta.

      42. - Sistema secondo la rivendicazione 26, caratterizzato dal fatto che detto convertitore è  
15 configurato per realizzare l'applicazione di due quantizzatori scalari con passo di quantizzazione scalato con costanti  $2/3$  e  $\sin(\Pi/3)$ , rispettivamente.

      43. - Sistema secondo la rivendicazione 42, caratterizzato dal fatto che detto convertitore è  
20 configurato per definire i punti di ricostruzione permessi in modo che questi formino un reticolo esagonale.

      44. - Sistema secondo la rivendicazione 26, caratterizzato dal fatto che detto convertitore è un  
25 codificatore configurato per:

- identificare un valore di nettezza del contorno in ciascuno di detti blocchi di pixel, suddividendo i contorni in più classi,

- attribuire valori di detto passo di  
30 quantizzazione (Q) differenziati in funzione delle classi così determinate.

      45. - Sistema secondo la rivendicazione 26 caratterizzato dal fatto che detto convertitore è configurato per determinare detto passo di  
35 quantizzazione (Q) secondo almeno una legge scelta nel

gruppo costituito da:

- $m \cdot E_Q + Q$  e
  - $Q = m \cdot (t^{\wedge} E_Q) + q$ , dove
- $m$  e  $q$  sono costanti, selettivamente determinate, e
- 5  $E_Q$  è un indice che identifica detta classe di contorno.

46. - Sistema secondo una qualsiasi delle rivendicazioni 29 a 32, caratterizzato dal fatto che detto convertitore è configurato per:

- 10 - suddividere detto livello di luminosità dei pixel nel blocco in più classi, e
- selezionare il valore di detto passo di quantizzazione ( $Q$ ) in modo differenziato in funzione delle classi così determinate.

15 47. - Sistema secondo una qualsiasi delle precedenti rivendicazioni 25 a 46, caratterizzato dal fatto che:

- detti segnali video in detto primo formato (10) sono segnali generati secondo un cosiddetto Bayer pattern, ordinabili in blocchi di dimensioni 4x2 aventi lo schema di componenti cromatiche:

riga 1 =  $G_1 R_1 G_2 R_2$ ,

riga 2 =  $B_1 G_3 B_2 G_4$ ,

e dal fatto che detto convertitore è configurato per

25 realizzare una quantizzazione vettoriale bidimensionale applicata alle coppie  $\langle R_1, R_2 \rangle$ ,  $\langle B_1, B_2 \rangle$ , insieme a  $\langle G_1, G_2 \rangle$ ,  $\langle G_3, G_4 \rangle$  oppure  $\langle G_1, G_3 \rangle$ ,  $\langle G_2, G_4 \rangle$ .

48.- Sistema secondo una qualsiasi delle rivendicazioni 26 a 47, caratterizzato dal fatto che

30 detti segnali video digitali in detto primo formato sono segnali digitali video in formato RGB e dal fatto che i segnali video digitali in detto secondo formato sono soggetti ad un cambiamento di coordinate ad uno spazio colore scelto dal gruppo costituito da YCbCr,

35 YUV, UIQ o YDbDr.

49. - Sistema secondo la rivendicazione 26 e la rivendicazione 48 caratterizzato dal fatto che detto quantizzatore vettoriale è applicato alle coppie adiacenti di pixel del piano di luminanza.

5        50. - Sistema secondo la rivendicazione 33, caratterizzato dal fatto che, in detti segnali video digitali in detto secondo formato, i piani di crominanza sono sottocampionati secondo un pattern a quinconce.

10       51. - Prodotto informatico direttamente caricabile nella memoria interna di un elaboratore numerico e comprendente porzioni di codice software per attuare il procedimento secondo una qualsiasi delle rivendicazioni 1 a 25 quando il prodotto è eseguito su un elaboratore  
15       numerico.

**RIASSUNTO**

Il sistema realizza la conversione di segnali video digitali organizzati in blocchi di pixel tra un primo formato (10, 30) ed un secondo formato (15, 25). Il  
5 secondo formato è un formato compresso tramite quantizzazione vettoriale (13, 27). La quantizzazione vettoriale è realizzata con l'applicazione ripetuta di un quantizzatore scalare applicato ai pixel di detti blocchi con passo di quantizzazione (Q) determinati in  
10 modo adattativo in funzione delle caratteristiche di nettezza e/o luminosità dei pixel.

(Figure 1 e 2)

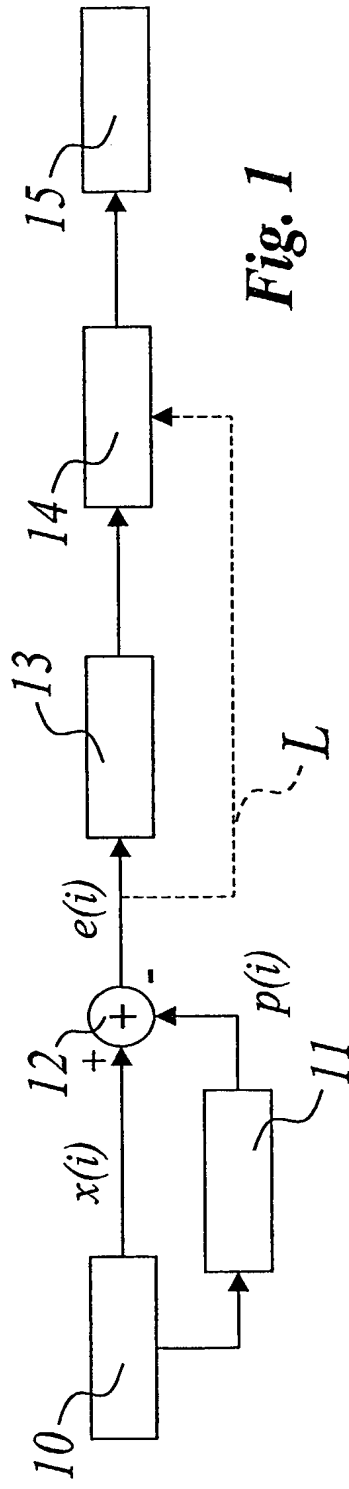


Fig. 1

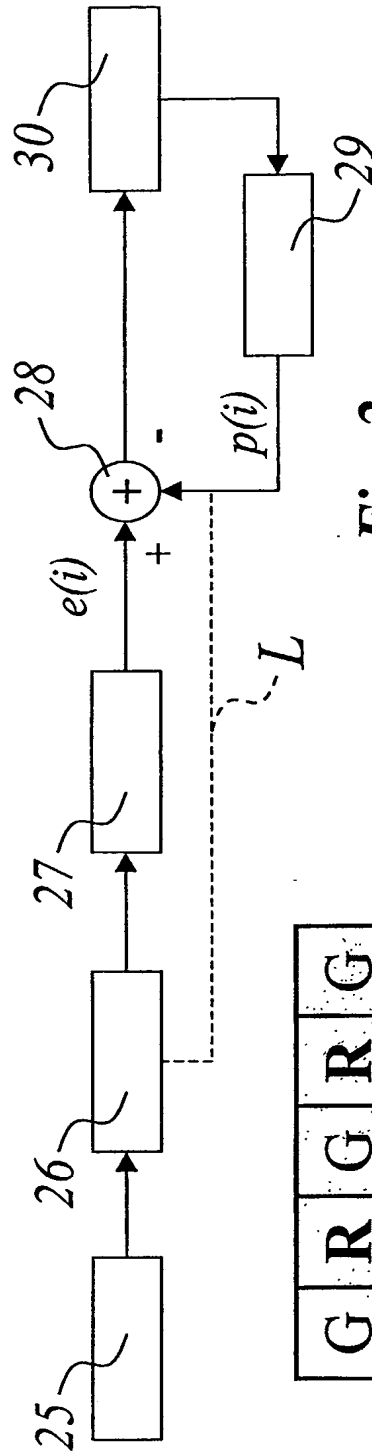
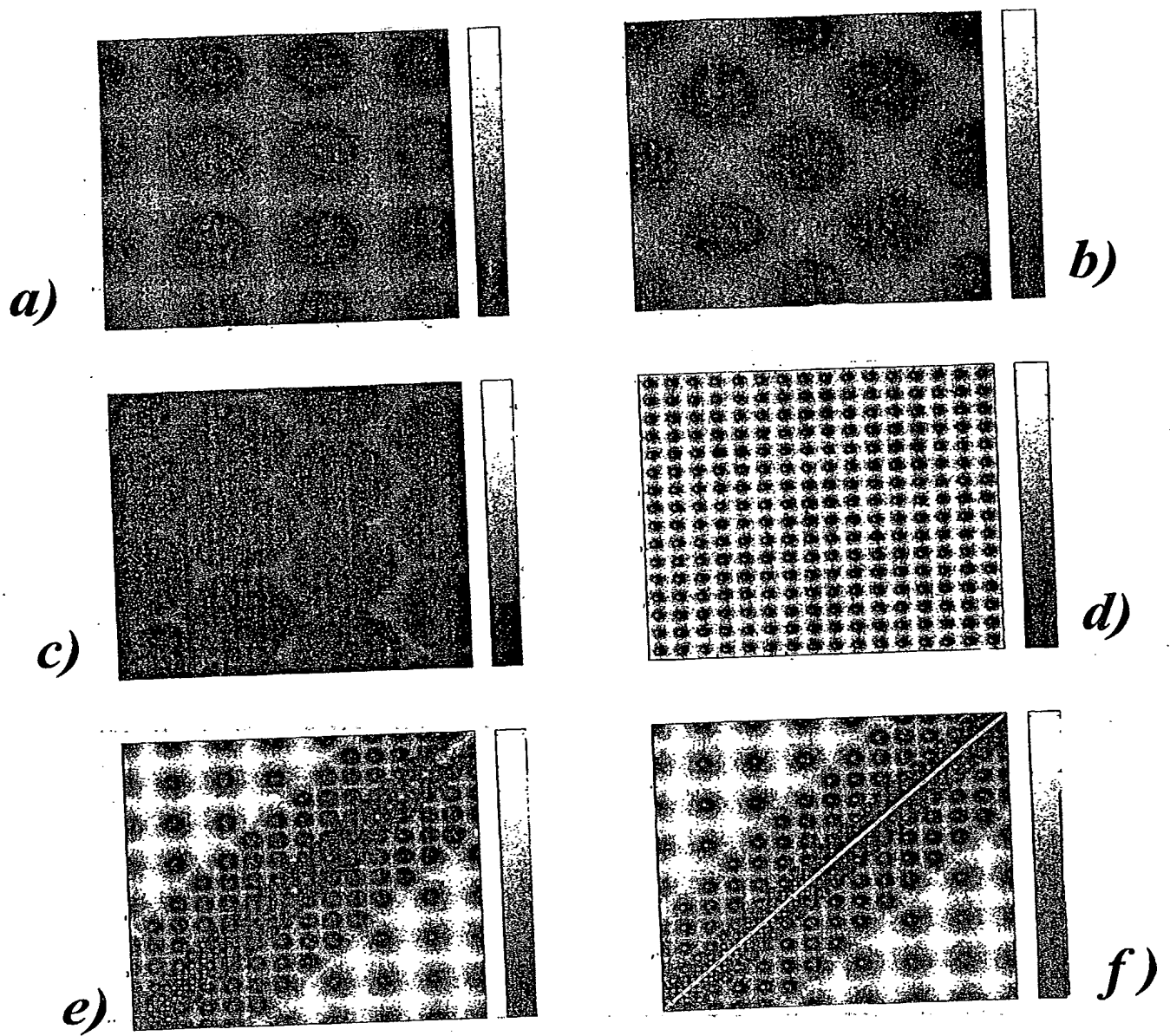


Fig. 2

Fig. 3 is a 5x5 grid of colored squares. The colors are arranged in a repeating pattern of rows and columns. The first row is G, R, G, R, G. The second row is B, G, B, G, B. The third row is G, R, G, R, G. The fourth row is B, G, B, G, B. The fifth row is G, R, G, R, G. The B squares are shaded with a stippled pattern.

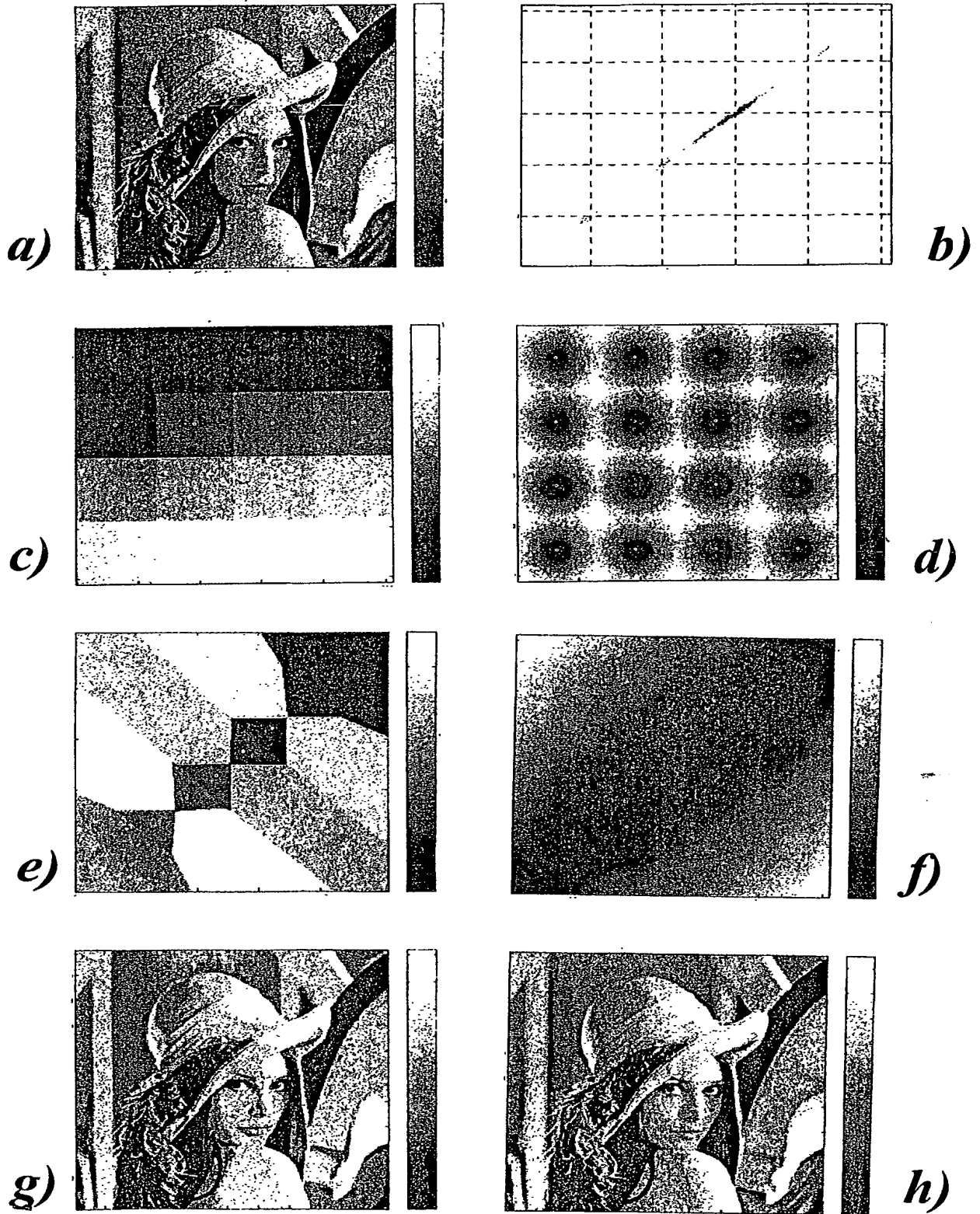
G	R	G	R	G
B	G	B	G	B
G	R	G	R	G
B	G	B	G	B
G	R	G	R	G

Fig. 3

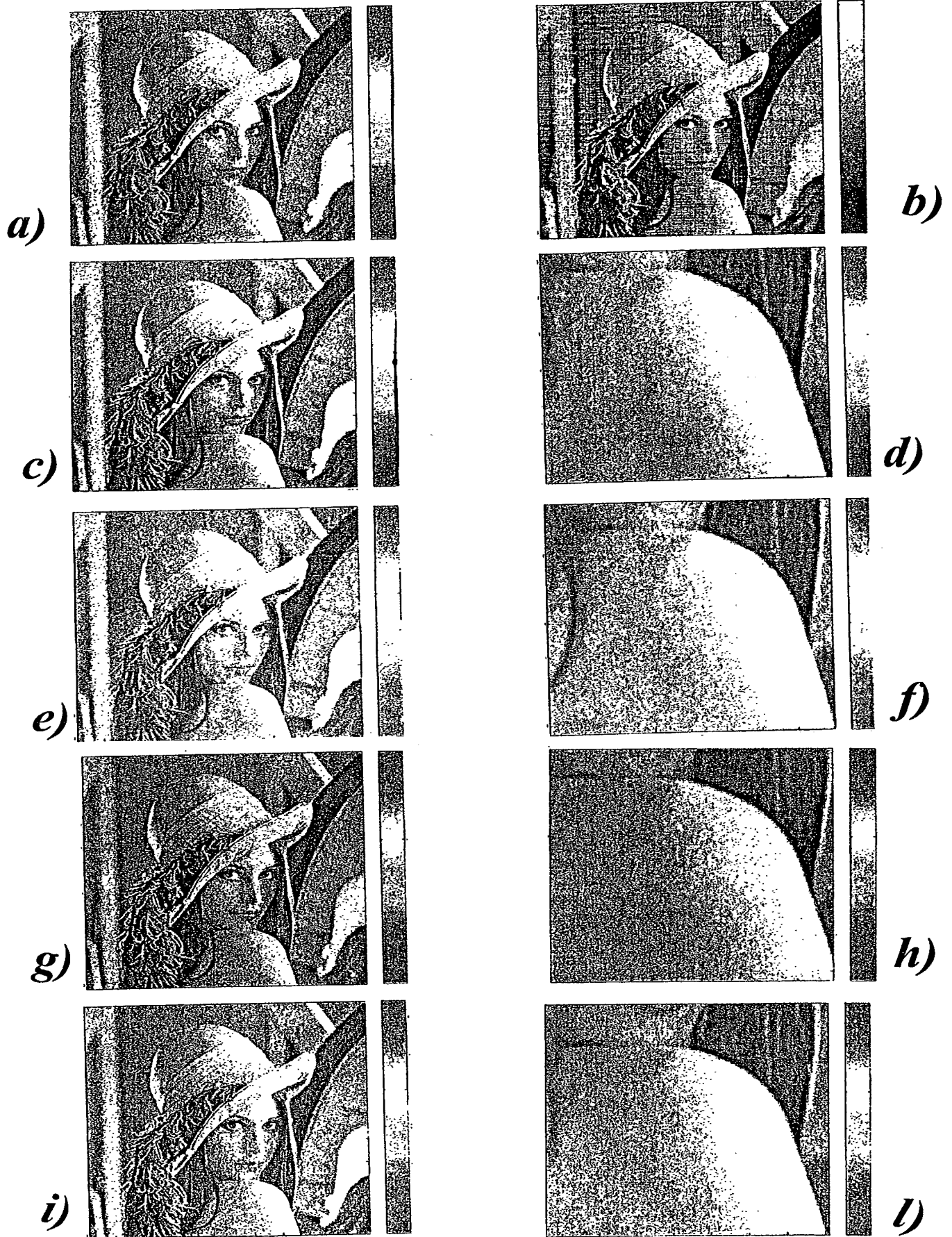


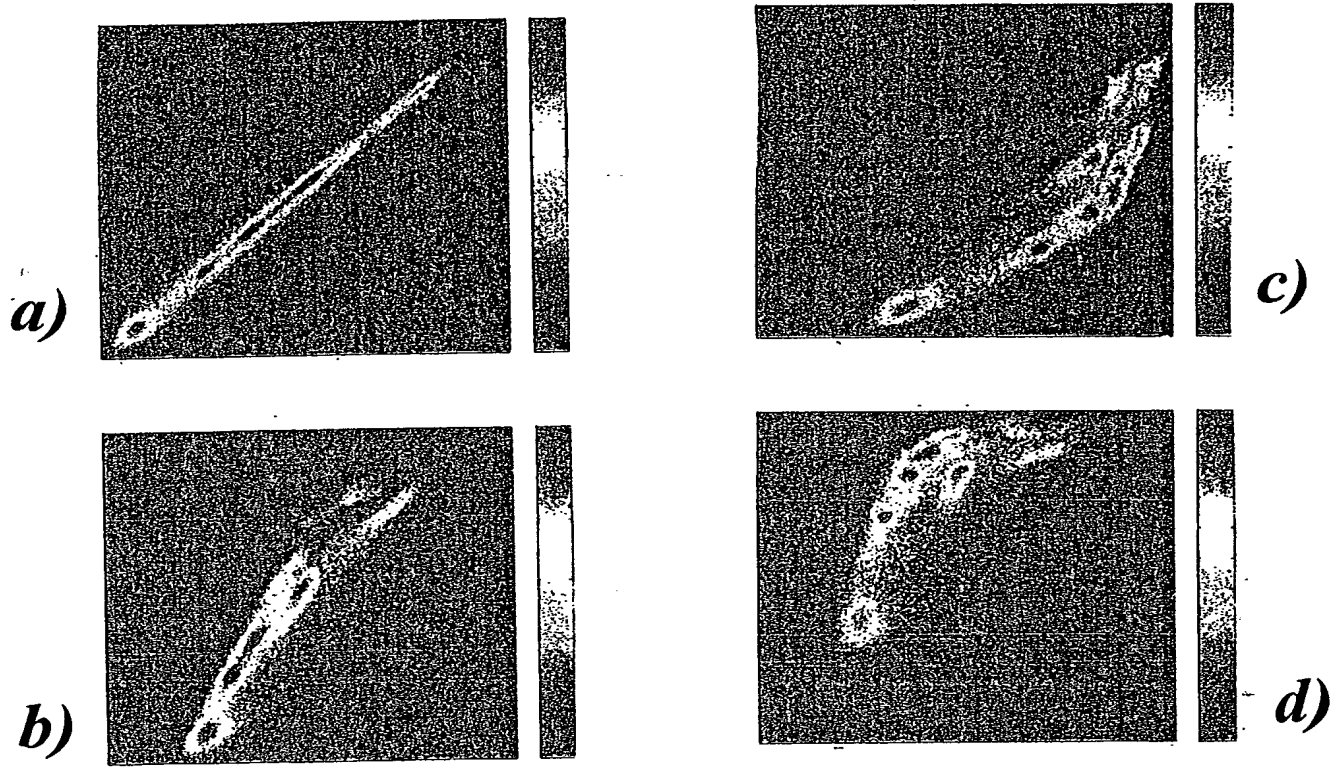
*Fig. 4*





*Fig. 5*

*Fig. 6*



**Fig. 7**

